

4w



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/057,161

01/24/2002

Liang He

42390P12178

1705

7590

01/19/2005

EXAMINER

ALBERTALLI, BRIAN LOUIS

James H. Salter  
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP  
7th Floor  
12400 Wilshire Boulevard  
Los Angeles, CA 90025

ART UNIT

PAPER NUMBER

2655

DATE MAILED: 01/19/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

**Office Action Summary**

Application No.

10/057,161

Applicant(s)

HE ET AL.

Examiner

Brian L Albertalli

Art Unit

2655

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☐ Responsive to communication(s) filed on \_\_\_\_.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-22 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-22 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date 10/10/02, 3/31/03, 11/17/03, 4/23/04
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: \_\_\_\_.

## DETAILED ACTION

### *Claim Objections*

1. Claims 1-22 are objected to for of the use of the term “DSR”. The term DSR is defined in the specification as “distributed speech recognition” (page 1, paragraph 2, line 1). However, specification also states that the DSR system was “standardized by the European Telecommunications Standards Institute (ETSI)” (page 2, paragraph 6, lines 1-2). It is not clear whether the term “DSR”, as used in the claims, is intended to encompass any distributed speech recognition system where speech features are sent from a client to a server for recognition processing, or only distributed speech recognition systems that conform to the specific standards of the ETSI. The term “DSR”, therefore, renders the claims indefinite (see 35 USC § 112, 2<sup>nd</sup> paragraph). For the purposes of examination, the term DSR has been interpreted herein as *any distributed speech recognition system where speech features are sent from a client to a server for recognition processing*.

2. Claims 2, 7, and 14 are objected to for of the use of the term “DSRML”. The acronym “DSRML” is not a well-known or common term in the art. The specification defines “DSRML” as “a specialized markup language based on conventional Extensible Markup Language (XML)” that “is defined and customized for the DSR application system”. Since the term “DSRML” is not a well-known or common term in the art, and given the definition given in the specification, it is unclear whether the term “DSRML” as cited in the claims is intended to encompass a variety of different specialized distributed

speech recognition languages, so long as they are based on XML, or whether DSRML is a specific (possibly proprietary) language for DSR applications. The term "DSRML", therefore, renders the claims indefinite (see 35 USC § 112, 2<sup>nd</sup> paragraph). For the purposes of examination, the term DSRML has been interpreted herein as *any language specialized for distributed speech recognition applications that conform to XML protocols*.

### ***Claim Rejections - 35 USC § 102***

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

4. Claims 1-4, 6-15, 17-19, and 21-22 are rejected under 35 U.S.C. 102(e) as being anticipated by Maes (U.S. Patent Application Publication 2002/0184373).

In regard to claim 1, Maes discloses a DSR system (Fig. 24) comprising:

a client (client 3000) to send connection requests, receive displayable content, and transmit speech feature data to a server (client 3000 includes a communications manager 3021 that provides synchronization protocols 3022 for synchronizing the GUI browser by processing event information (sending requests and receiving displayable

content), page 32, paragraph 322, lines 12-15 and paragraph 324, lines 1-4; and transports DSR encoded data to the DSR decoders 3012, page 32, paragraph 323);

a gateway coupled between the client and the server to support data communication between the client and the server (EDGE server 3025 includes a gateway 3026 to convert client and server requests across different network protocols, page 32, paragraph 325); and

a server (3004) to receive the speech feature data, perform speech recognition on the speech feature data, and transmit displayable content to the client (DSR decoder 3012 decodes the DSR encoded data and conversational engines 3011 perform speech recognition, page 32, paragraph 323 and paragraph 326, lines 1-3; multi-modal shell 3013 updates the views displayed through GUI browser 3001, page 31, paragraph 321, line 21 through page 32, 1<sup>st</sup> column, lines 1-3).

In regard to claim 2, Maes discloses:

client wrapper API to interface with a DSR client browser (wrapper 3003, page 31, paragraph 321, lines 12-16);

a DSR frame constructor coupled to the client wrapper API to construct DSR frames;

a DSR payload wrapper coupled to the DSR frame constructor to construct DSR payload packets from the DSR frames (communications manager 3021 supports the voice coding and transport protocols of 3023, page 32, paragraph 323; the DSR transport protocol layer implements RTP and RTCP to packet the speech data, page

21, paragraph 210, lines 1-4; the RTP packages frame pairs (FP) into DSR payloads, see Fig. 29(a) and page 11, paragraph 118, lines 16-23); and

a DSRML client transceiver to receive displayable content and to send an initial connection request to the server (multi-modal shell 3013 updates the views displayed through GUI browser 3001, page 31, paragraph 321, line 21 through page 32, 1<sup>st</sup> column, lines 1-3; communications manager 3021 includes synchronization protocols 3022 that remotely control the conversational speech engines using the SERCP protocol, page 32, paragraph 324; the SERCP protocol being a DSR protocol based on XML, page 26, paragraph 263, lines 1-3).

In regard to claim 3, Maes discloses a client transmission/recognition adapter to adjust transmission control conditions of the DSR payload wrapper and to control flag bits needed for speech recognition according to transmission/recognition parameters; and

said DSR payload wrapper to add flag bits to the DSR payload packets (the RTP protocol generates a header that contains various flag bits indicating the size and type of the payload frames, see Fig. 7, and page 11, paragraph 118, lines 1-3; and appends RTP header to the payload frame pairs, see Fig. 29(a) and page 11, paragraph 119, lines 1-8).

In regard to claim 4, Maes discloses a client protocol stack having a TCP module supporting TCP protocol and an IP module supporting IP protocol (edge server 3025

communicates with the client through the TCP/IP protocol, page 32, paragraph 325, lines 1-5).

In regard to claim 6, Maes discloses a feature compressor coupled to the client wrapper API and the DSR frame constructor to compress speech feature data (page 9, paragraph 101, lines 1-5).

In regard to claim 7 Maes discloses:

a DSR payload de-wrapper to separate DSR speech feature data from transmission/recognition parameters;

a DSR frame extractor coupled to the DSR payload de-wrapper to extract DSR frames (DSR decoder 3012 decodes the encoded voice data to provide feature data to speech recognizer 3011, page 32, paragraph 323);

a server wrapper API coupled to the DSR frame extractor to interface with a DSR server browser (wrapper 3008, page 31, paragraph 321, lines 12-16); and

a DSRML server transceiver to send displayable content and to receive an initial connection request from the client (multi-modal shell 3013 updates the views displayed through GUI browser 3001, page 31, paragraph 321, line 21 through page 32, 1<sup>st</sup> column, lines 1-3; communications manager 3021 includes synchronization protocols 3022 that remotely control the conversational speech engines using the SERCP protocol, page 32, paragraph 324; the SERCP protocol being a DSR protocol based on XML, page 26, paragraph 263, lines 1-3).

In regard to claim 8, Maes discloses a server stack having a UDP module to support UDP protocol, the server further including:

an RTP receiver to receive DSR payload packets using RTP through UDP/IP protocol stacks and extracting DSR payload from the DSR payload packets; and

a server transmission/recognition adapter coupled to the DSR payload de-wrapper and the DSR frame extractor to control frame extraction according to transmission parameters and flag bits for speech recognition (voice coding and transport protocols 3023, page 32, paragraph 323; implement RTP and RTCP to packet and control the encoded speech data over the network, page 21, paragraph 210, lines 1-4; over a UDP layer (see Fig. 20, 2007), page 20, paragraph 203, lines 17-18; the server must inherently include the necessary extractors and de-wrappers for decoding the speech feature data).

In regard to claim 9, Maes discloses a feature compressor on the client side compresses the feature data before transmission to the server (Fig. 2b, 214).

In regard to claims 10 and 11, Maes discloses the gateway supports both wired and wireless data communication (3G wireless network and H.323 wired network, page 20, paragraph 203, lines 5-11).



In regard to claim 12, Maes discloses a web server coupled to the server via a network (content server 3014, page 31, paragraph 320, lines 9-11).

In regard to claim 13, Maes discloses a front-end engine for reducing noise and to extract the speed feature data (page 22, paragraph 228).

In regard to claim 14, Maes discloses the displayable content is represented as a DSRML document (multi-modal shell 3013 updates the views displayed through GUI browser 3001, page 31, paragraph 321, line 21 through page 32, 1<sup>st</sup> column, lines 1-3; communications manager 3021 includes synchronization protocols 3022 that remotely control the conversational speech engines using the SERCP protocol, page 32, paragraph 324; the SERCP protocol being a DSR protocol based on XML, page 26, paragraph 263, lines 1-3).

In regard to claims 15 and 19, Maes discloses a method and machine readable medium having executable code stored thereon to perform the method, comprising:

- receiving input speech data;

- extracting speech features from the input speech data;

- packaging the speech features into DSR frames in a DSR frame format (Fig. 2a, MFCC's are extracted from input speech, page 9, paragraph 100; MFCC features are a framed format)

collecting DSR frames to form a DSR payload (page 10, paragraph 111, lines 1-4); and

transmitting the DSR payload to a server for speech recognition processing (page 10, paragraph 110, lines 11-13).

In regard to claims 17 and 21, Maes discloses a method and machine readable medium having executable code stored thereon to perform the method, comprising:

- receiving a DSR payload packet;
- de-wrapping DSR payload from the DSR payload packet and separating DSR speech feature data from transmission/recognition parameters;
- extracting DSR frames from the DSR payload (the DSR decoder must inherently have the necessary extractors and de-wrappers to recover the speech feature data);
- extracting speech feature data from the DSR frames (Fig. 2b, 213 and page 9, paragraph 104, lines 7-10); and
- sending the speech feature data to a speech recognition engine and for recognition (Fig. 2b, 212, page 9, paragraph 104, lines 5-7).

In regard to claims 18 and 22, Maes discloses de-compressing the speech feature data (decompression module 214, page 9, paragraph 104, lines 7-9).

***Claim Rejections - 35 USC § 103***

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claim 5 is rejected under 35 U.S.C. 103(a) as being unpatentable over Maes.

Maes discloses that the voice coding and transport protocols (3023, page 32, paragraph 323) implement RTP and RTCP to packet and control the encoded speech data over the network (page 21, paragraph 210, lines 1-4) over a UDP layer (see Fig. 20, 2007, page 20, paragraph 203, lines 17-18).

Maes does not specifically disclose that the RTP sender includes a buffer to store packets that have been sent, but not acknowledged by the server, and retransmitting the RTP packets until they are acknowledged by the server.

Official notice is taken that it is notoriously well known and recognized in the art to include a buffer in an RTP implementation over UDP, since UDP does not include any assurance that packets have reached their destination. Including a buffer to store the RTP until they are acknowledged as being received by an RTCP packet ensures that every packet reaches the server. This would be especially important in a DSR system, since any loss of speech feature packets would result in extreme recognition errors.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Maes to include an RTP buffer, in order to ensure that every packet reached the server, thereby ensuring the best recognition results possible.

7. Claims 16 and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Maes, as applied to claims 15 and 19, above, in view of Allman et al. (*Increasing TCP's Initial Window*), in further view of Nishimura et al. (U.S. Patent 6,754,200), and further in view of Mathis et al. (*TCP Selective Acknowledgement Options*).

Maes discloses passing the DSR payload to a transport protocol stack composed of TCP and IP (client and server communicate through the TCP/IP protocol (page 32, paragraph 325).

Maes does not disclose:

increasing a TCP initial window;

Allman et al. disclose increasing a TCP initial window reduce the transmission time for connections transmitting only a small amount of data (page 3, section 3, second paragraph).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Maes to increase the TCP initial window, in order to reduce the transmission time of the speech features (which are a small amount of data), so that the overall latency of the system was reduced, providing a faster response time to a user's spoken commands.

Neither Maes nor Allman et al. disclose:

adopting no slow-start restart;

Nishimura et al. (U.S. Patent 6,754,200) disclose a system that adopts a no slow-start restart when a discarded packet is not congestion related (column 9, lines 42-46).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Maes and Allman et al. to adopt a no slow-start restart so that there would be no waiting time for retransmission if a packet was lost, as taught by Nishimura et al. (column 9, lines 56-57). Again, this would reduce the overall latency of the system, providing a faster response time to a user's spoken commands.

Neither Maes nor Allman et al. nor Nishimura et al. disclose applying a TCP SACK.

Mathis et al. disclose multiple packet losses have a catastrophic on TCP throughput because it causes the client (sender) to retransmit segments which have already been correctly received (page 2, first paragraph, lines 1-7). SACK corrects that by ensuring that the client (sender) only needs to retransmit packets that have actually been lost (page 2, second paragraph).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Maes, Allman et al., and Nishimura et al. to apply TCP SACK to increase the overall throughput of the system in situations where there was multiple packet losses, again reducing the overall latency of the system and providing a faster response time to a user's spoken commands.

**Conclusion**

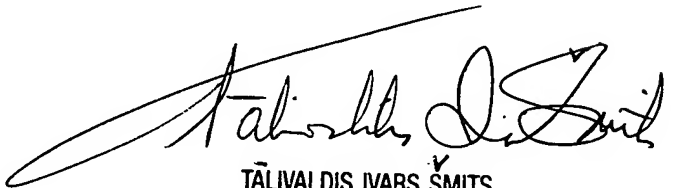
8. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Coyner et al. (*Distributed Speech Recognition Services*) and Zhang et al. (*The Study on Distributed Speech Recognition Systems*) both disclose the typical DSR architecture. Maes et al. (U.S. Patent 6,801,604) disclose a system for generating DSR payloads and the associated transmission and decoding of those DSR payloads. Kushner et al. (U.S. Patent Application Publication 2002/0147579) disclose a DSR system with a particular front-end configuration. He et al. (U.S. Patent Application Publication 2002/0184197) disclose a DSR system for use over a wireless network. Boloker et al. (U.S. Patent Application Publication 2002/0194388) disclose a system for building multi-modal browsers for a DSR system. Bergman et al. (U.S. Patent Application Publication 2003/0161298) disclose a wireless DSR system over WAP gateway.

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brian L Albertalli whose telephone number is (703) 305-1817. The examiner can normally be reached on Mon - Fri, 8:00 AM - 5:30 PM, every second Fri off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 305-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

BLA 1/14/05



TĀIVALDIS IVARS SMITS  
PRIMARY EXAMINER